ELSEVIER Linear Algebra and its Applications 275–276 (1998) 161–177

LINEAR ALGEBRA AND ITS APPLICATIONS

Differentiable structure of the set of controllable $(A,B)^t$-invariant subspaces [1]

J. Ferrer *, F. Puerta, X. Puerta

Departament de Matemàtica Aplicada I, E.T.S. Enginyers Industrials de Barcelona, Universitat Politècnica de Catalunya, Diagonal 647, E-08028 Barcelona, Spain

Received 25 November 1996; accepted 1 September 1997

Submitted by V. Mehrmann

Abstract

Given $(A,B)^t \in \operatorname{Hom}(\mathbb{C}^n, \mathbb{C}^{n+m})$ observable, we prove that the set of $(A,B)^t$-invariant subspaces having a fixed Brunovsky–Kronecker structure is a connected manifold, and we compute its dimension. Also, we include some applications of these results. © 1998 Elsevier Science Inc. All rights reserved.

1. Introduction

Given linear maps $A^t : \mathbb{C}^n \to \mathbb{C}^n$ and $B^t : \mathbb{C}^n \to \mathbb{C}^m$, a subspace $W \subset \mathbb{C}^n$ is called $(A,B)^t$-invariant if $A^t(W \cap \operatorname{Ker} B^t) \subset W$ (see [1], Theorem 6.6.1, p. 207). It is equivalent to say that $W^\perp$ is $(\overline{A},\overline{B})$-invariant (see [1], p. 207), so that the analysis of the $(A,B)^t$-invariant subspaces is dual to the one of the (A,B)-invariant subspaces. This kind of subspaces play a key role in solving many problems in the Linear Systems Theory. See for example [2,1], where the algebraic structure of such subspaces is studied and some interesting applications to the Disturbance Decoupling Problem and the Output Stabilization Problem are presented.

* Corresponding author.

[1] Partially supported by DGICYT no. PB94-1365-C03-03.

0024-3795/98/$19.00
PII: S0024-3795(97)10062-3

We will deal with $(A,B)^{\mathrm{t}}$-invariant subspaces because of their geometric (coordinate free) study is more convenient. Indeed, in [3] (Section 5) we show that the geometric study of the maps $(A,B)^{\mathrm{t}}: \mathbb{C}^n \to \mathbb{C}^{n+m}$ is just the one of the linear maps $f: \mathscr{Y} \to \mathfrak{X}$, where $\mathscr{Y}$ is a subspace of $\mathfrak{X}$: it is sufficient to consider $(A,B)^{\mathrm{t}}$ as the matrix of f in some basis of $\mathfrak{X}$ adapted to $\mathscr{Y}$. In particular, taking $\mathfrak{X} = \mathbb{C}^{n+m}$ and $\mathscr{Y} = \mathbb{C}^n \equiv \mathbb{C}^n \times \{0\} \subset \mathfrak{X}$, so that $\operatorname{Ker} B^{\mathrm{t}} = f^{-1}(\mathscr{Y})$, we adapt the above definition as follows ([4], Definition 3.1; see also [5]): a subspace $\mathscr{S}$ of $\mathscr{Y}$ is called f-invariant if $f(\mathscr{S}) \cap \mathscr{Y} \subset \mathscr{S}$. Of course, if $\mathscr{Y} = \mathfrak{X}$ this definition is the usual one for endomorphisms.

For the case of f being an endomorphism, the structure of the set of invariant subspaces has been studied, for example, in [1,6]. In particular, the last one studies its differentiable structure: it need not be a manifold; however, it does the subset consisting of the invariant subspaces with a fixed cyclic structure. Our aim is to obtain an analogous result for $(A,B)^{\mathrm{t}}$-invariant subspaces, when the pair $(A,B)^{\mathrm{t}}$ is observable.

In the free coordinate language, this condition of observability can be reformulated as follows (see [3], Section 3.3): the only subspace $\mathscr{S} \subset \mathscr{Y}$ such that $f(\mathscr{S}) \subset \mathscr{S}$ in $\mathscr{S} = \{0\}$. Then, we will say that f is observable. In what follows, we assume that $f: \mathscr{Y} \to \mathfrak{X}$ is observable.

For $f: \mathscr{Y} \to \mathfrak{X}$ a linear map defined on a subspace, there is a canonical form, the Brunovsky–Kronecker-form, which generalizes the Jordan form for an endomorphism. (In all the paper, BK-(...) means Brunovsky–Kronecker-(...)). Then, in an analogous way to [6], one can consider the set of d-dimensional f-invariant subspaces $\mathscr{S}$ of $\mathscr{Y}$ such that the restriction $\hat{f}: \mathscr{S} \to \mathfrak{X}$ has a given BK-form

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix},$$

compatible with that of $f: \mathscr{Y} \to \mathfrak{X}$. Denote this set by $\operatorname{Inv}(f;(M,F))$. We show that it is a connected manifold (4.5 and 5.2), and we compute its dimension (Theorem. 5.3).

As an application, we prove (5.1) the existence of global differentiable BK-bases for a family of subspaces in $\operatorname{Inv}(f;(M,F))$, differentiabily parametrized over a contractible manifold.

The structure of the paper is as follows. Section 2 contains the notation used in the sequel, and the definition (Definition 2.1) of the set $\operatorname{Inv}(f;(M,F))$.

In Section 3 we study the BK-bases of the subspaces in $\operatorname{Inv}(f;(M,F))$, or rather their coordinates in a given BK-basis of f: if we write them as the columns of a matrix, they are just the solutions of the matrix system (a)–(c) in Theorem 3.1; moreover, we explicit them in 3.8.

By means of this characterization, in Section 4 we construct a natural differentiable structure of $\operatorname{Inv}(f;(M,F))$ as a homogeneous space (Theorem 4.5).

In Section 5 we prove that $\mathrm{Inv}(f;(M,F))$ is connected (5.2) and we compute its dimension (5.3). Also, it contains the refered application to global BK-bases (5.1), and some examples of "extreme" dimension (5.4).

$\mathcal{M}_{p,q}$ will denote the set of complex matrices having p-rows and q-columns, and $\mathcal{M}^*_{p,q}$ the ones having maximal rank. If $p=q$, we will write simply $\mathcal{M}_p$ and $\mathcal{M}^*_p$, respectively. The latter, with the group structure of matrix multiplication, is the linear group $\mathrm{Gl}(\mathbb{C}^p)$.

For any $\mathbb{C}$-vector space $\mathscr{Z}$, $\mathrm{Gr}_d(\mathscr{Z})$ will denote the grassman manifold of d-dimensional subspaces of $\mathscr{Z}$.

For the notions about bundles appearing in Sections 4 and 5, the reader may find a simplified survey in Section 3 of [7].

2. The set Inv (f; (M, F)) of f-invariant subspaces having the same BK-matrix

We fix an $(n+m)$-dimensional vector space $\mathfrak{X}$ over the complex numbers $\mathbb{C}$, an n-dimensional linear subspace $\mathscr{Y} \subset \mathfrak{X}$, and a linear map $f:\mathscr{Y} \to \mathfrak{X}$ defined on it. We assume that f is observable, it is to say, that the only subspace $\mathscr{S} \subset \mathscr{Y}$ such that $f(\mathscr{S}) \subset \mathscr{S}$ is $\mathscr{S} = \{0\}$.

We recall (see [1] or [3]) that there exist so-called BK-bases of $f:\mathscr{Y} \to \mathfrak{X}$ of the form $B=(B_{\mathrm{O}},B_{\mathrm{E}},B_{\mathrm{A}})$, where B is a basis of $\mathfrak{X}$ and B_{O} is a basis of $\mathscr{Y}$ such that: B_{O} is formed by so called BK-chains $w_i, f(w_i), \ldots, f^{k_i-1}(w_i)$, $1 \leqslant i \leqslant r$, $k_1 \geqslant \cdots \geqslant k_r$; B_{E} is the family $f^{k_1}(w_1), \ldots, f^{k_r}(w_r)$, formed by the ends of the BK-chains; B_{A} is arbitrary.

The integers $k_1, \ldots, k_r$ do not depend on the choice of the BK-basis, and are called the BK-indices of f.

Then, in any bases B_{O} and B of this kind, the matrix of f is

$$\begin{pmatrix} N \\ E \\ 0 \end{pmatrix} \in \mathcal{M}_{n+m,n},$$

where: $N = \mathrm{diag}\{N_1, \ldots, N_r\}$, each N_i being the standard nilpotent k_i-square matrix,

$$N_i = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

and $E = \mathrm{diag}\{E_1, \ldots, E_r\}$, each $E_i = (0\ 0 \ldots 0\ 1)$ being a k_i-row matrix. This is called the BK-matrix of f.

We fix one of these BK-bases of $f:\mathscr{Y} \to \mathfrak{X}$. For the sequel we identify the vectors of $\mathfrak{X}$ (respectively $\mathscr{Y}$) with the $(n+m)$-columns (respectively, n-columns) matrices of their coordinates in this basis. In the same sense, we identify $\mathrm{Gr}_d(\mathscr{Y})$ with $\mathrm{Gr}_d(\mathbb{C}^n)$, for any d.

On the other hand, we recall that a subspace $\mathscr{S} \subset \mathscr{Y}$ is called f-invariant if $f(S) \cap \mathscr{Y} \subset S$ (see the introduction). Our aim is to study the differentiable structure of the set of f-invariant subspaces $\mathscr{S} \subset \mathscr{Y}$ such that the BK-matrix of the restriction $\hat{f}: \mathscr{S} \to \mathfrak{X}$ be a given one. In order to do that, we fix a family of BK-indices $h_1 \geqslant \cdots \geqslant h_s$ and a BK-matrix

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix} \in \mathscr{M}_{n+m,d},$$

where, in an analogous way to

$$\begin{pmatrix} N \\ E \\ 0 \end{pmatrix}, \quad M = \mathrm{diag}\{M_1, \ldots, M_s\},$$

each M_i being the standard nilpotent h_i-square matrix, etc.

Definition 2.1. With the above notation, we denote by $\mathrm{Inv}(f; (M,F)) \equiv \mathrm{Inv}(f; (h_1, \ldots, h_s))$ the set of subspaces $\mathscr{S} \subset \mathscr{Y}$ such that: $\mathscr{S}$ is f-invariant, and the BK-matrix of the restriction $\hat{f}: \mathscr{S} \to \mathfrak{X}$ is

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix},$$

or equivalently, the BK-indices of $\hat{f}$ are $h_1, \ldots, h_s$.

Remark 2.2. We assume that the conditions that guarantee that $\mathrm{Inv}(f; (M,F)) \neq \emptyset$ are satisfied (see [5,4]): $s \leqslant r$, and $h_i \leqslant k_i$ for $i = 1, 2, \ldots$

3. Characterization of the BK-bases of an f-invariant subspace

In 3.1 we will characterize the BK-bases of the subspaces in $\mathrm{Inv}(f; (M,F))$, and we will derive their explicit form in 3.8. In fact, this special form of the BK-bases of the subspaces in $\mathrm{Inv}(f; (M,F))$ can be obtained directly. However, their characterization by means of the conditions (a)–(c) in 3.1 will be useful in the sequel. Let Φ be the map

$$\Phi: \mathscr{M}^*_{n,d} \quad \to \quad \mathrm{Gr}_d(\mathscr{Y})$$

such that: if $X \in \mathscr{M}^*_{n,d}$, $\Phi(X)$ is the subspace $\mathscr{S} \in \mathrm{Gr}_d(\mathscr{Y})$ spanned by the columns of X. For simplicity, we say that X is a basis of $\mathscr{S} = \Phi(X)$.

In addition, we say that X is a BK-basis of $\mathscr{S} = \Phi(X)$ if it can be extended to a BK-basis of $\hat{f}: \mathscr{S} \to \mathfrak{X}$.

Theorem 3.1. *With the above notation:*
1. Let $\mathscr{S} \in \mathrm{Gr}_d(\mathscr{Y})$ such that

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix}$$

is the BK-matrix of the restriction $\hat{f}: \mathscr{S} \to \mathfrak{X}$. If X is a BK-matrix of $\mathscr{S}$, then (a) $NX = XM + NXF^{t}F$, (b) $EX = EXF^{t}F$. Moreover, $\mathscr{S}$ is f-invariant if and only if (c) EXF^{t} has maximal rank.
2. Conversely, let $X \in \mathscr{M}^{*}_{n,d}$. If X satisfies the conditions (a)–(c) above, then $\mathscr{S} = \Phi(X) \in \mathrm{Inv}(f; (M, F))$ and X is a BK-matrix of it.

Proof. 1. Firstly, let us see that: if X is a BK-basis of $\mathscr{S}$, then there is some matrix R such that the matrix

$$Q =: \left(\begin{array}{c|c|c} X & NXF^{t} & \\ \hline 0 & EXF^{t} & R \\ \hline 0 & 0 & \end{array}\right)$$

is invertible. In fact, we will take Q a BK-basis of $\hat{f}: \mathscr{S} \to \mathfrak{X}$ of the form $\hat{B} = (\hat{B}_{\mathrm{O}}, \hat{B}_{\mathrm{E}}, \hat{B}_{\mathrm{A}})$, in an analogous way to B in Section 2. Due to the hypothesis, there is some basis $\hat{B}$ of this kind such that the columns of X are the vectors in $\hat{B}_{\mathrm{O}}$, which we denote by $(\hat{w}_i, \dots, \hat{f}^{h_i-1}(\hat{w}_i))$, $1 \leqslant i \leqslant s$. Thus, the vectors $\hat{f}^{h_i-1}(\hat{w}_i)$, $1 \leqslant i \leqslant s$, are the columns of XF^{t}; hence, the vectors in $\hat{B}_{\mathrm{E}}$ must be the columns of

$$\begin{pmatrix} N \\ E \\ 0 \end{pmatrix} XF^{t}.$$

Finally, we take the columns of R as the vectors in $\hat{B}_{\mathrm{A}}$. Next, since Q is a BK-basis of $\hat{f}: \mathscr{S} \to \mathfrak{X}$, the matrix of $\hat{f}$ in this basis must be

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix},$$

so that:

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix} = Q^{-1} \begin{pmatrix} N \\ E \\ 0 \end{pmatrix} X$$

and conditions (a) and (b) follow immediately.

For condition (c), recall that the columns of

$$\begin{pmatrix} N \\ E \\ 0 \end{pmatrix} XF^{t}$$

are the coordinates of the vectors $f^{h_1}(\hat{w}_1), \ldots, f^{h_s}(\hat{w}_s)$ in the basis $B = (B_{\mathrm{O}}, B_{\mathrm{E}}, B_{\mathrm{A}})$. In particular, the columns of EXF^{t} are their B_{E}-coordinates. So, condition (c) is equivalent to $\{\alpha_1 \hat{f}^{h_1}(\hat{w}_1) + \cdots + \alpha_s \hat{f}^{h_s}(\hat{w}_s), \alpha_i \in \mathbb{C}\} \cap \mathscr{Y} = \{0\}$. This is to say, $f(S) \cap \mathscr{Y} \subset S$.

2. Conversely, if $X \in \mathscr{M}^{*}_{n,d}$ satisfies the condition (c), then there is R such that

$$Q = \left(\begin{array}{c|c|c} X & NXF^{t} & \\ \hline 0 & EXF^{t} & R \\ \hline 0 & 0 & \end{array} \right)$$

is invertible. If, in addition, X satisfies (a) and (b), then

$$Q^{-1} \begin{pmatrix} N \\ E \\ 0 \end{pmatrix} X = \begin{pmatrix} M \\ F \\ 0 \end{pmatrix}$$

Taking into account that the columns of Q are a basis of $\mathfrak{X}$ such that the first d vectors are a basis of $\mathscr{S} = \Phi(X)$, we conclude that Q is a BK-basis of $\hat{f}: \mathscr{S} \to \mathfrak{X}$, where $\mathscr{S} = \Phi(X)$, and the BK-matrix of $\hat{f}$ is

$$\begin{pmatrix} M \\ F \\ 0 \end{pmatrix}.$$

Moreover, by 3.1(1), condition (c) implies that $\mathscr{S}$ is f-invariant. $\square$

This result motivates the following:

Definition 3.2. We denote by $\mathscr{M}((N,E);(M,F)) \equiv \mathscr{M}((k_1,\ldots,k_r);(h_1,\ldots,h_s))$, or simply by $\mathscr{M}$ if no confusion is possible, the set of matrices $X \in \mathscr{M}^{*}_{n,d}$ which satisfy conditions (a)–(c) in Theorem 3.1.

Obviously, $\mathcal{M}$ is a submanifold of $\mathcal{M}^*_{n,d}$. In fact, an open subset of a linear subvariety of $\mathcal{M}_{n,d}$. With this notation, Theorem 3.1 can be reformulated as follows.

Corollary 3.3. *With the above notation, we have*

$$\Phi(\mathcal{M}((N,E);(M,F))) = \operatorname{Inv}(f;(M,F)).$$

We are going to obtain the explicit form of the matrices in $\mathcal{M}((N,E);(M,F))$. Firstly, we will solve the equations (a) and (b) in 3.1. Next, in 3.7, we will see that the condition that X has maximal rank follows from (a)–(c) in 3.1, so that it can be dropped in Definition 3.2

Proposition 3.4. *Let $X(k,h)$ denote the following matrix of $\mathcal{M}_{k,h}$:*

$$X(k,h) = \begin{pmatrix} x_1 & \dots & \dots & 0 \\ x_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ x_{k-h} & \ddots & \ddots & x_1 \\ x_{k-h+1} & \ddots & \ddots & x_2 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_{k-h} \\ 0 & \dots & \dots & x_{k-h+1} \end{pmatrix} \quad \text{if} \quad k \geqslant h$$

$$X(k,h) = 0 \qquad \text{if} \quad k < h$$

the non-zero coefficients being arbitrary. Then, the solutions of the equations (a) and (b) of 3.1 are the matrices of the form

$$X = \begin{pmatrix} X(k_1,h_1) & \dots & X(k_1,h_s) \\ \vdots & \dots & \vdots \\ X(k_r,h_1) & \dots & X(k_r,h_s) \end{pmatrix} \in \mathcal{M}_{n,d},$$

where each block $X(k_i,h_j)$ is as above. We will denote by $x_1(k_i,h_j),\dots,x_{k_i-h_j+1}(k_i,h_j)$ the independent variables which appears in the (first column of the) block $X(k_i,h_j)$, if $k_i \geqslant h_j$.

Proof. Decompose X into blocks

$$X = \begin{pmatrix} X_{11} & \dots & X_{1s} \\ \vdots & \dots & \vdots \\ X_{r1} & \dots & X_{rs} \end{pmatrix},$$

where $X_{ij} \in \mathscr{M}_{k_i,h_j}$. Then X is solution of (a) and (b) of 3.1 if and only if each one of its blocks X_{ij} satisfies

$$N_iX_{ij} = X_{ij}M_j + N_iX_{ij}F_j^{t}F_j,$$
$$E_iX_{ij} = E_iX_{ij}F_j^{t}F_j,$$

$1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$. So, it is sufficient to solve (a) and (b) of 3.1 when N and M are standard nilpotent blocks. In this case, if $X = (x_{ij})$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$ it is easy to check that the entries x_{ij} of X are the solutions of the following system:

$$\begin{aligned}
&x_{12} = \cdots = x_{1h} = 0\\
&x_{11} = x_{22}, x_{12} = x_{23}, \ldots, x_{1h-1} = x_{2h}\\
&\vdots\\
&x_{k-1,1} = x_{k2}, x_{k-1,2} = x_{k3}, \ldots, x_{k-1,h-1} = x_{kh}\\
&x_{k1} = x_{k2} = \cdots = x_{kk-1} = 0
\end{aligned}$$

Hence, $X = X(k, h)$, and the proposition follows. $\square$

Remark 3.5. 1. The form of the solutions of (a) and (b) in 3.1 described in the above proposition is motivated by the fact (easily checked) that X has the above form if and only if the columns of X are chains of vectors of the form

$$v_1, f(v_1), \ldots, f^{h_1-1}(v_1), \ldots, v_s, f(v_s), \ldots, f^{h_s-1}v(_s).$$

2. If X is a matrix as in 3.4, then we have the following expression for the matrix EXF^{t}: EXF^{t} is the submatrix of X formed by the entries in the right-bottom corner of each block $X(k_i, h_j)$; it is to say, $EXF^{t} = (x_j^i)$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$, where $x_j^i = x_{k_i-h_j+1}(k_i, h_j)$ if $k_i \geqslant h_j$, or 0 otherwise.

Example 3.6. The matrices $\mathscr{M}((4, 2, 1); (3, 1))$ are those of form

$$\left(\begin{array}{ccc|c}
x_1(4,3) & 0 & 0 & x_1(4,1)\\
x_2(4,3) & x_1(4,3) & 0 & x_2(4,1)\\
0 & x_2(4,3) & x_1(4,3) & x_3(4,1)\\
0 & 0 & x_2(4,3) & x_4(4,1)\\
\hline
0 & 0 & 0 & x_1(2,1)\\
0 & 0 & 0 & x_2(2,1)\\
0 & 0 & 0 & x_1(1,1)
\end{array}\right)$$

such that they have maximal rank and the submatrices

$$\begin{pmatrix} x_2(4,3) & x_4(4,1) \\ 0 & x_2(2,1) \\ 0 & x_1(1,1) \end{pmatrix}$$

have also maximal rank.

In fact the condition that X has maximal rank can be dropped.

Proposition 3.7. *Let $X \in \mathcal{M}_{n,d}$ be a solution of (a) and (b) of* 3.1 *as in* 3.4. *Then if EXF^{t} has maximal rank, X has also maximal rank.*

Proof. Let $v_1, f(v_1), \ldots, f^{h_1-1}(v_1), \ldots, v_s, f(v_s), \ldots, f^{h_s-1}(v_s)$ be the columns of X and $\mathcal{S}$ the subspace of $\mathfrak{X}$ spanned by this set of vectors. We have to show that they are linearly independent. As we have seen in the proof of 3.1, if $EXF^{\mathrm{t}} = (x_j^i)$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$, then

$$f^{h_j}(v_j) = \sum_{i=1}^{r} x_j^i f^{k_i}(w_i) + y_j,$$

where $y_j \in \mathcal{Y}$, $1 \leqslant j \leqslant s$. Since EXF^{t} has maximal rank, the vectors $f^{h_1}(v_1), \ldots, f^{h_s}(v_s)$ are linearly independent. Again as we have seen in the proof of 3.1, taking into account that $\mathcal{S} \subset \mathcal{Y}$ we have

$$\operatorname{span}\{f^{h_1}(v_1), \ldots, f^{h_s}(v_s)\} \cap \mathcal{S} = \{0\}.$$

Moreover, f being injective by observability, we have the following chain of implications (l.i. means linearly independent)

$$\begin{aligned} &f^{h_1}(v_1), \ldots, f^{h_s}(v_s) \text{ l.i. and } \operatorname{span}\{f^{h_1}(v_1), \ldots, f^{h_s}(v_s)\} \cap \mathcal{S} = \{0\} \Rightarrow \\ &f^{h_1-1}(v_1), \ldots, f^{h_s-1}(v_s) \text{ l.i. and } \operatorname{span}\{f^{h_1-1}(v_1), \ldots, f^{h_s-1}(v_s)\} \cap f^{-1}(\mathcal{S}) \\ &= \{0\} \Rightarrow \\ &f^{h_1-2}(v_1), \ldots, f^{h_s-2}(v_s) \text{ l.i. and } \operatorname{span}\{f^{h_1-2}(v_1), \ldots, f^{h_s-2}(v_s)\} \cap f^{-2}(\mathcal{S}) \\ &= \{0\} \Rightarrow \\ &\ldots \end{aligned}$$

Now, suppose that

$$\sum_{j=1}^{s} \sum_{i=0}^{h_j-1} \lambda_i^j f^i(v_j) = 0$$

then

$$\sum_{j=1}^{s} \sum_{i=0}^{h_j-1} \lambda_i^j f^{i+1}(v_j) = 0$$

and we conclude that $\lambda_{h_1-1}^1 = \cdots = \lambda_{h_s-1}^s = 0$. In a similar way we prove that $\lambda_i^j = 0$ for $1 \leqslant i \leqslant h_j - 1$, $1 \leqslant j \leqslant s$. $\square$

Summarizing 3.4 and 3.7 we have the following.

Corollary 3.8. *The set $\mathcal{M}((N,E);(M,F))$ is formed by the matrices $X \in \mathcal{M}_{n,d}$ such that (i) X has the form described in* 3.4, (ii) *the submatrix EXF^t (see* 3.5 (2)) *has maximal rank.*

Next, we will relate the different BK-bases of a given subspace in $\mathrm{Inv}(f;(M,F))$. In general, the map Φ is not injective. In fact, we have that $\Phi(X) = \Phi(X')$ if and only if $X' = XT$ for some $T \in \mathrm{Gl}(\mathbb{C}^d)$. If only matrices in $\mathcal{M}$ are considered (see 3.3), we have the following.

Proposition 3.9. *Let $X, X' \in \mathcal{M}$. Then, $\Phi(X) = \Phi(X')$ if and only if there is $T \in Gl(\mathbb{C}^d)$ such that $X' = XT$, and:*
(a′) $MT = TM + MTF^tF$,
(b′) $FT = FTF^tF$.

Proof. Let $X \in \mathcal{M}$, $T \in \mathrm{Gl}(\mathbb{C}^d)$ satisfying (a′) and (b′), and $X' = XT$. Then,

$$NX' = NXT = XMT + NXF^tFT = XTM + XMTF^tF$$
$$+ NXF^tFTF^tF = X'M + (XM + NXF^tF)TF^tF = X'M$$
$$+ NXTF^tF = X'M + NX'F^tF,$$
$$EX' = EXT = EXF^tFT = EXF^tFTF^tF = EXTF^tF = EX'F^tF.$$

Moreover, by (2) of 3.1, $\mathcal{S} = \Phi(X)$ is f-invariant. As $\Phi(X) = \Phi(X')$, it follows from (1c) of 3.1 that $EX'F^t$ has maximal rank.

Conversely, assume that $X, X' \in \mathcal{M}$, and $X' = XT$ with $T \in \mathrm{Gl}(\mathbb{C}^d)$. By applying (1b) of 3.1 for $X' = XT$ and for X successively we have

$$EXT = EXTF^tF = EXF^tFTF^tF.$$

On the other hand, if we multiply by T, condition (b) for X, we obtain

$$EXT = EXF^tFT.$$

Hence

$$EXF^tFTF^tF = EXF^tFT.$$

Since EXF^t has full column rank, it follows:

$$FTF^tF = FT.$$

For (a′), in an analogous way, and using the last equality, we have

$$NXT = XTM + NXTF^tF = XTM + XMTF^tF$$
$$+ NXF^tFTF^tF = XTM + XMTF^tF + NXF^tFT.$$

On the other hand, multiplying (1a) of 3.1 by T, we obtain

$$NXT = XMT + NXF^{\mathrm{t}}FT.$$

Hence

$$XTM + XMTF^{\mathrm{t}}F = XMT.$$

Since X has maximal rank, if follows:

$$TM + MTF^{\mathrm{t}}F = MT. \qquad \square$$

4. The differentiable structure of Inv(f;(M,F))

The above proposition suggests the following definition.

Definition 4.1. We denote by $\mathscr{G}(M,F)$, or simply by $\mathscr{G}$ if no confusion is possible, the set of matrices $T \in \mathrm{Gl}(\mathbb{C}^d)$ which satisfy conditions (a′) and (b′) in Proposition 3.9.

Lemma 4.2. *With the above notation, if $T \in \mathscr{G}$, then (c′) FTF^{t} has maximal rank. In fact, $(FTF^{\mathrm{t}})^{-1} = FT^{-1}F^{\mathrm{t}}$.*

Proof. By applying (b′), we have

$$FTF^{\mathrm{t}}FT^{-1}F^{\mathrm{t}} = FTT^{-1}F^{\mathrm{t}} = FF^{\mathrm{t}} = I_s. \qquad \square$$

Remark 4.3. Because of the last lemma, we can identify $\mathscr{G}$ with $\mathscr{M}((M,F);(M,F))$. However, we are mainly interested in the group structure of $\mathscr{G}$.

Lemma 4.4. *With the above notation:*
1. *$\mathscr{G}$ is a subgroup of $\mathrm{Gl}(\mathbb{C}^d)$*
2. *$\mathscr{G}$ acts freely on $\mathscr{M}$ on the right by matrix multiplication.*

Proof. 1. The proof that $T, T' \in \mathscr{G}$ implies $TT' \in \mathscr{G}$ is analogous to the first part of 3.9. It is straightforward that: $MF^{\mathrm{t}}F = 0$, $FF^{\mathrm{t}} = I_s$. So that $I_s \in \mathscr{G}$. Let us see that $T^{-1} \in \mathscr{G}$ if $T \in \mathscr{G}$. From (b′) we have

$$(FTF^{\mathrm{t}})^{-1}(FT)T^{-1} = FT^{-1}$$

and the left member is $FT^{-1}F^{\mathrm{t}}F$ because of 4.2. Finally, by multiplying by T^{-1} on both sides of (a′) we have

$$MT^{-1} = T^{-1}M - T^{-1}MTF^{\mathrm{t}}FT^{-1}.$$

Next, we apply successively (b′) for T^{-1} and (a′) for T

$$\begin{aligned} -T^{-1}MTF^{\mathrm{t}}FT^{-1} &= -T^{-1}MTF^{\mathrm{t}}FT^{-1}F^{\mathrm{t}}F \\ &= -T^{-1}(MT \\ -TM)T^{-1}F^{\mathrm{t}}F &= -T^{-1}MF^{\mathrm{t}}F + MT^{-1}F^{\mathrm{t}}F. \end{aligned}$$

2. We proved in 3.9 that $XT \in \mathcal{M}$ if $X \in \mathcal{M}$ and $T \in \mathcal{G}$. The action is free if $XT = X$ implies $T = I$. And this is so because X has maximal rank. □

Since $\mathcal{G}$ acts on $\mathcal{M}$, we can consider the orbit $X\mathcal{G}$ of an element $X \in \mathcal{M}$ which is the set $\{XT; T \in \mathcal{G}\}$. Now a natural differentiable structure in $\mathrm{Inv}(f;(M,F))$ can be defined by means of the following theorem.

Theorem 4.5. *Let $\mathcal{M}/\mathcal{G}$ be the set of orbits under the action* 4.4, *and $\tilde{\Phi}$ the map induced on it by Φ. Then:*

1. $\tilde{\Phi}: \mathcal{M}/\mathcal{G} \to$ $\mathrm{Inv}(f;(M,F))$ is a bijection.

2. The orbit space $\mathcal{M}/\mathcal{G}$ has a differentiable structure such that the natural projection $\pi: \mathcal{M} \to \mathcal{M}/\mathcal{G}$ is a submersion. In fact, $\pi: \mathcal{M} \to \mathcal{M}/\mathcal{G}$ is a principal bundle with structural group $\mathcal{G}$.

3. $\dim(\mathcal{M}/\mathcal{G}) = \dim \mathcal{M} - \dim \mathcal{G}$.

Proof. Assertion (1) follows from the above lemma and Corollary 3.3. For (2), it is sufficient to prove (see, for example, [8], Theorem (2.9.10), p. 82 and Remark p. 83) that:

(a) $\delta = \{(X, X') \in \mathcal{M} \times \mathcal{M}: X' = XT$ for some $T \in \mathcal{G}\}$ is closed in $\mathcal{M} \times \mathcal{M}$.

(b) The map $\gamma: \mathcal{M} \times \mathcal{G} \to \delta$ defined by $\gamma(X,T) = (X, XT)$, is a homeomorphism.

To prove (a), we consider

$$\Delta = \{(X, X') \in \mathcal{M}^*_{n,d} \times \mathcal{M}^*_{n,d}: X' = XT \quad \text{for some} \quad T \in Gl(\mathbb{C}^d)\}$$

Clearly, $\delta = \Delta \cap (\mathcal{M} \times \mathcal{M})$, and Δ is closed in $\mathcal{M}^*_{n,d} \times \mathcal{M}^*_{n,d}$. (Notice that $\mathcal{M}^*_{n,d}/\mathrm{Gl}(\mathbb{C}^d) = \mathrm{Gr}_d(\mathbb{C}^n)$.)

Concerning (b), it is obvious that γ is onto, and the injectivity follows immediately from the above lemma. Finally, notice that γ is the restriction of the homeomorphism

$$\Gamma: \mathcal{M}^*_{n,d} \times \mathrm{Gl}_d(C) \quad \to \quad \Delta, \Gamma(X,T) = (X, XT).$$

Finally, (3) follows from (2) taking into account the local triviality property of the principal bundles. □

Remark 4.6. In the conditions of the above theorem, it is known that the following properties are satisfied.

1. Each orbit $X\mathcal{G} = \{XT, T \in \mathcal{G}\}$ is a closed submanifold of $\mathcal{M}$, diffeomorphic to $\mathcal{G}$.

2. For any differentiable manifold $\mathcal{N}$, a map $\psi\colon \mathcal{M}/\mathcal{G} \to \mathcal{N}$ is smooth if and only if $\psi \circ \pi$ is smooth. In particular, $\tilde{\Phi}\colon \mathcal{M}/\mathcal{G} \to \mathrm{Gr}_d(\mathcal{Y})$ is smooth.

In fact, notice that (b) in the proof of the above theorem implies that the action of $\mathcal{G}$ is proper. Then (1) follows from (4.1.22) of [9], p. 265. And (2) is an easy consequence of the existence of local sections of the principal bundle π (see [9], p. 264).

Remark 4.7. One of the referees has pointed out that in the case of invariant subspaces of an endomorphism f, Shayman ([6]) did not only show that $\mathrm{Inv}(f)$ has the structure of a manifold but also there is a variety structure, so that it would be interesting to know if it is also true in the case of $\mathrm{Inv}(f;(M,F))$.

5. Further properties and applications

As a first application of 4.5, we study the existence of global differentiable BK-bases for a differentiably parametrized family of subspaces in $\mathrm{Inv}(f;(M,F))$.

We consider a manifold $\mathcal{W}$, and a family $\mathcal{S}(t)$, $t \in \mathcal{W}$, of subspaces in $\mathrm{Inv}(f;(M,F))$, differentiably parametrized over $\mathcal{W}$. It can simply be represented as a differentiable map $\mathcal{S}(t)\colon \mathcal{W} \to \mathrm{Inv}(f;(M,F))$. We know that, for each $t \in \mathcal{W}$, there are some BK-bases $X_t \in \mathcal{M}((N,E);(M,F))$ of $\mathcal{S}(t)$. We show that, locally, it is possible to choose one of these BK-bases $X(t)$ for each t, in such a way that it depends differentiably of t. Moreover, this result globalizes to all $\mathcal{W}$ if it is contractible.

Proposition 5.1. *Let $\mathcal{W}$ be a manifold, and $\mathcal{S}(t)\colon \mathcal{W} \to \mathrm{Inv}(f;(M,F))$ a differentiable map.*

1. For each $t_0 \in \mathcal{W}$ there exist an open neighborhood $\mathcal{U}$ of t_0 in $\mathcal{W}$, and a differentiable map $X(t)\colon \mathcal{U} \to \mathcal{M}((N,E);(M,F))$ such that $X(t)$ is a BK-basis of $\mathcal{S}(t)$ for all $t \in \mathcal{U}$.

2. If in addition $\mathcal{W}$ is contractible, the statement in part 1 *holds for $\mathcal{U} = \mathcal{W}$.*

Proof. 1. It is sufficient to take a local section $\sigma\colon \mathcal{U} \to \mathcal{M}$ of the submersion π (see 4.5), and to define $X(t) = \sigma(\mathcal{S}(t))$ for all $t \in \mathcal{U}$.

2. According to Corollary 3 (and Theorem 2) of [10], there exists a differentiable basis $v_1(t),\dots,v_d(t),\dots,v_{n+m}(t)$ of $\mathfrak{X}$ adapted to $\mathcal{S}(t)$ for all $t \in \mathcal{W}$. Then, for each $t \in \mathcal{W}$, let

$$\begin{pmatrix} \hat{A}(t) \\ \hat{C}(t) \end{pmatrix}$$

be the matrix of the restriction of f to $\mathscr{S}(t)$, in the corresponding basis. Clearly, it is a differentiable family of pairs of matrices, parametrized over $\mathscr{W}$. And by hypothesis it has constant BK-indices $h_1, \ldots, h_s$, so that we can apply Theorem 6 of [10] and the proof is concluded. An alternative proof, analogous to the one of (4.2.3) in [7], follows from the existence of a global section of the $\mathscr{G}$-principal bundle $\mathscr{S}^*(\pi): \mathscr{S}^*(\mathscr{M}) \to \mathscr{W}$ induced by pullback under $\mathscr{S}(t)$ from $\pi: \mathscr{M} \to \mathrm{Inv}(f; (M, F))$. □

Further topological and differentiable properties can be derived from 3.8 and 4.5.

Proposition 5.2. *The manifold* $\mathrm{Inv}(f; (M, F))$ *is connected.*

Proof. Clearly (see 4.5), it is sufficient to prove that $\mathscr{M}$ is connected. We consider the description of $\mathscr{M}$ in 3.8. If we denote by $\bar{\mathscr{M}}$ the linear subspace of $\mathscr{M}_{n,d}$ formed by the matrices described in 3.4, we have

$$\mathscr{M} = \{X \in \bar{\mathscr{M}} : EXF^{\mathrm{t}} \quad \text{has maximal rank}\}.$$

According to 3.5(2), this restriction only involves the variables $x_{k_i-h_j+1}(k_i, h_j)$, $k_i \geqslant h_j$. Hence, for any other variable $x_v(k_i, h_j)$ we have

$$\mathscr{M} = (\mathscr{M} \cap \{x_v(k_i, h_j) = 0\}) \times \mathbb{C},$$

where we have identified $\mathbb{C}$ with the $x_v(k_i, h_j)$-axis. Therefore, $\mathscr{M}$ is connected if and only if this is true for the intersection $\mathscr{M} \cap \{x_v(k_i, h_j) = 0\}$. By repeating this process, we reduce our study to the subset

$$\mathscr{M}_0 = \mathscr{M} \cap \{x_1(k_i, h_j) = \cdots = x_{k_i-h_j}(k_i, h_j) = 0, \quad k_i > h_j\}.$$

Next, notice that a matrix $X \in \mathscr{M}_0$ has at maximum one non-zero variable in each block $X(k_i, h_j)$, repeated along the bottom diagonal. We denote by x^i_j these unique variables, as in 3.5(2): $x^i_j \equiv x_{k_i-h_j-1}(k_i, h_j)$, if $k_i \geqslant h_j$; $x^i_j = 0$ otherwise. Therefore, we can identify a matrix X of $\mathscr{M}_0$ with its submatrix $EXF^{\mathrm{t}} = (x^i_j)$ formed by the right-bottom entry of each block. That is to say, we can identify $\mathscr{M}_0$ with the set $\mathscr{M}'^*$ defined as follows:

$$\mathscr{M}' = \{X' = (x^i_j) \in \mathscr{M}_{r,s} : x^i_j = 0 \text{ if } k_i < h_j\},$$
$$\mathscr{M}'^* = \{X' \in \mathscr{M}' : X' \text{ has maximal rank } (= s)\}.$$

Hence, it is sufficient to prove that $\mathscr{M}'^*$ is connected.

In order to that, notice that the matrices $X' \in \mathscr{M}'$ are block-echelon of the form

$$X' = \begin{pmatrix} X'_{11} & X'_{12} & X'_{13} & \dots \\ 0 & X'_{22} & X'_{23} & \dots \\ 0 & 0 & X'_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix},$$

where all the null blocks are placed under the diagonal $x_1^1, \dots, x_s^s$ (because of $k_i \geqslant h_i$). Then, the subset

$$\mathscr{M}'^{**} = \{X' \in \mathscr{M}' : X'_{11}, X'_{22}, \dots, \quad \text{have maximal rank}\}$$

is dense in $\mathscr{M}'^{*}$ (in fact, any $X' \in \mathscr{M}'$ is adherent to $\mathscr{M}'^{**}$, because the blocks X'_{vv} became of maximal rank by means of a generic perturbation), so that it is sufficient to prove that $\mathscr{M}'^{**}$ is connected.

For that, we reduce, as above, the study to the subset

$$\mathscr{M}_0'^{**} = \mathscr{M}'^{**} \cap \{X'_{v\eta} = 0, \quad v \neq \eta\} = \{X' \in \mathscr{M}'^{**} : X' = \operatorname{diag}\{X'_{11}, X'_{22}, \dots\}\}$$

so that $\mathscr{M}_0'^{**}$ is the cartesian product of sets of the form $\mathscr{M}^*_{p,q}$. And this kind of sets are connected because they are the complementary subsets (in $\mathscr{M}_{p,q}$) of the intersection of

$$\binom{p}{q}$$

hypersurfaces, each one defined by the annulation of a q-minor. □

Proposition 5.3. *With the notation in Section 2*

$$\begin{aligned}\dim \operatorname{Inv}(f; (M,F)) &= \sum_{\substack{1 \leqslant i \leqslant r \\ 1 \leqslant j \leqslant s}} \sup\{k_i - h_j + 1, 0\} - \sum_{1 \leqslant i,j \leqslant s} \sup\{h_i - h_j + 1, 0\} \\ &= \sum_{1 \leqslant i \leqslant s} (k_i - h_i)(s - i + 1) + \sum_{\substack{1 \leqslant j < i \leqslant s \\ h_j < h_i}} (k_i - h_j) \\ &\quad + \sum_{\substack{1 \leqslant j < i \leqslant s \\ h_j > h_i}} \sup\{k_i - h_j + 1, 0\} + \sum_{\substack{s < i \leqslant r \\ 1 \leqslant j \leqslant s}} \sup\{k_i - h_j + 1, 0\}.\end{aligned}$$

Proof. Because of 3.8, we have

$$\dim \mathscr{M}((N,E); (M,F)) = \sum_{\substack{1 \leqslant i \leqslant r \\ 1 \leqslant j \leqslant s}} \sup\{k_i - h_j + 1, 0\}.$$

In an analogous way

$$\dim \mathscr{M}((M,F); (M,F)) = \sum_{1 \leqslant i,j \leqslant s} \sup\{h_i - h_j + 1, 0\}.$$

Then, by 4.5

$$\begin{aligned}\dim\ \operatorname{Inv}(f;(M,F)) =& \dim\mathscr{M}((N,E);(M,F)) \\ & - \dim\mathscr{M}((M,F);(M,F)),\end{aligned}$$

so that we have proved the first equality.

The second one is obtained by considering separately the following four subsets of indices: $1 \leqslant i \leqslant j \leqslant s$; $1 \leqslant j < i \leqslant s$, $h_j = h_i$; $1 \leqslant j < i \leqslant s$, $h_j > h_i$; $s < i \leqslant r$, $1 \leqslant j \leqslant s$. $\square$

As a final application, we consider some special cases.

Example 5.4. By applying 5.3, we have

(a) $\dim \operatorname{Inv}((k_1, \dots, k_r); (h_1, \dots, h_s)) = 0$ only in the following case:

$$k_s > k_{s+1}, \quad h_j = k_j \quad \text{for all} \quad 1 \leqslant j \leqslant s.$$

(b) $\dim\ \operatorname{Inv}((k_1, \dots, k_r); (h_1, \dots, h_s)) = 1$ only in the following two cases:

(i) $k_s = k_{s+1} > k_{s+2}, \quad h_j = k_j \quad \text{for all} \ 1 \leqslant j \leqslant s-1.$

(ii) $k_{s-1} > k_s > 1 + k_{s+1}, \quad h_s = k_s - 1, \quad h_j = k_j \ \textit{for all} \quad 1 \leqslant j \leqslant s-1.$

(c) Let $\mathscr{Y} \subset \mathfrak{X}$ and $f \colon \mathscr{Y} \to \mathfrak{X}$ such that: $\dim \mathscr{Y} = kr$, $k_1 = \cdots = k_r = k$. Let consider the set of r-dimensional f-invariant subspaces of $\mathscr{Y}$ such that: $s = r$, $h_1 = \cdots = h_s = 1$. We have

$$\dim \operatorname{Inv}((k, \dots, k); (1, \dots, 1)) = r^2(k-1).$$

Notice that it is the dimension of the grassman manifold of all r-dimensional subspaces of $\mathscr{Y}$.

Acknowledgements

We are grateful to the referees for their valuable suggestions and their careful review.

References

[1] I. Gohberg, P. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, Wiley, New York, 1986.

[2] W.M. Wonham, Linear Multivariable Control: A Geometric Approach, Springer, New York, 1979.

[3] J. Ferrer, F. Puerta, Similarity of non-everywhere defined linear maps, Linear Algebra Appl. 168 (1992) 27–55.

[4] A. Compta, J. Ferrer, On $(A,B)^t$-invariant subspaces having extendible Brunovsky bases, Linear Algebra Appl. 255 (1997) 185–201.

[5] I. Baragaña, I. Zaballa, Block similarity invariants of restrictions to (A, B)-invariants subspaces, Linear Algebra Appl. 220 (1995) 31–62.

[6] M.A. Shayman, On the variety of invariant subspaces of a finite-dimensional linear operator, Trans. AMS 274 (2) (1982) 721–747.

[7] J. Ferrer, M.I. García, F. Puerta, Differentiable families of subspaces, Linear Algebra Appl. 199 (1994) 229–252.

[8] V.S. Varadarajan, Lie Groups, Lie Algebras and their Representation, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[9] R. Abraham, J.E. Marsden, Foundations of Mechanics, Benjamin/Cummings, Menlo Park, CA, 1978.

[10] J. Ferrer, F. Puerta, Global block-similarity and pole assignment of class C^p, SIMAX (To appear).